ELSEVIER

Gene 294 (2002) 249–257

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Characterization of IS*999*, an unstable genetic element in *Mycobacterium avium*

Jean-Pierre Laurent, Shawn Faske, Gerard A. Cangelosi*

Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109, USA

Received 7 February 2002; received in revised form 7 June 2002; accepted 19 June 2002
Received by A.M. Campbell

Abstract

An IS3-family insertion element, IS*999*, was identified in the opportunistic pathogen *Mycobacterium avium*. The 1347 bp element has 29 bp inverted repeats and two overlapping open reading frames coding for putative transposases. It was detected in the genomes of ten of 12 *M. avium* isolates examined. Copy numbers ranged from four to 16. IS*999* is less stable than IS*1245*, the most commonly-used marker for typing *M. avium* isolates. Among 60 colonies picked from a single patient isolate, there were two distinct IS*1245* restriction fragment length polymorphism banding patterns compared to eight distinct IS*999* patterns (five in one IS*1245* group and three in the other). In view of its instability, we asked whether transposition of IS*999* might have phenotypic consequences. Nucleotide sequence analysis of insertion sites in four isolates revealed 16 putative structural genes that were variably disrupted by IS*999*. Insertions into *hdhA*, a gene that codes for a putative short chain alcohol dehydrogenase, were distributed non-randomly between colony type variants, consistent with phenotypic consequences that exert selective pressure. These observations illustrate the genetic heterogeneity that can exist within populations of *M. avium* that appear to be homogeneous by IS*1245* analysis. IS*999* may be a useful marker for tracking, at the sub-strain level, the rapid genetic drift that *M. avium* isolates undergo in nature and in the laboratory.

Keywords: Insertion element; Morphotypic switching; Transposition; Restriction-modification; Short chain alcohol dehydrogenase; MAC (*Mycobacterium avium* complex)

1. Introduction

Bacterial insertion sequences (ISs) have been associated with antibiotic resistance, gene acquisition, pathogenicity, and modification of metabolic pathways (Campbell, 1981; Hernandez-Alles et al., 1999; Mahillon and Chandler, 1998). They have been shown to participate in chromosome rearrangements, contributing to the evolution of bacterial genomes (Gaffney and Lessie, 1987; Sreevatsan et al., 1997; Sampson et al., 1999; Fang et al., 1999). ISs are common in *Mycobacterium* species. The genome of *Mycobacterium tuberculosis* strain H37Rv has 56 loci with homology to ISs (Gordon et al., 1999). Because of its stability and its widespread occurrence in *M. tuberculosis* isolates, IS*6110* has become the standard marker for molecular epidemiological analysis of this pathogen by restriction fragment length polymorphism (RFLP) (van Embden et al., 1993). More recently, IS*1245* was adapted as the standard RFLP marker for *Mycobacterium avium*, the most significant of the environmental mycobacteria that opportunistically infect humans (van Soolingen et al., 1998). Although less stable than IS*6110*, IS*1245* is more suitable for this purpose than other mobile elements in *M. avium* (Bauer et al., 1999).

The stability of insertion elements in *M. avium* may be limited in part by their genomic environment. As a species *M. avium* is a more heterogeneous than *M. tuberculosis* and it is known for its genomic and phenotypic fluidity. Polyclonal infections of humans are common (Arbeit et al., 1993; Slutsky et al., 1994; Wallace et al., 1998). Even within populations of cells that appear to be monoclonal, diagnosis and treatment are complicated by colony type variation, the basis for which remains poorly understood (Matsiota-Bernard et al., 2000; Arbeit et al., 1993; Belisle and Brennan, 1994). The rough colony type of *M. avium* was recently shown to result from genetic deletions involving

Abbreviations: COG, clusters of orthologous groups; CR, Congo Red; IS-APD, insertion sequence-directed amplified polymorphic DNA; MAG, Middlebrook 7H9-albumin-glycerol medium; ORF, open reading frame; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; RO, red opaque colony type; WO, white opaque colony type

* Corresponding author. Tel.: +1-206-284-8846, ext. 313; fax: +1-206-284-0313.
E-mail address: gcang@sbri.org (G.A. Cangelosi).

0378-1119/02/$ - see front matter
PII: S0378-1119(02)00797-7

IS*1601* in some isolates (Eckstein et al., 2000). Recently, we described a new phenotypic variation in *M. avium* (Cangelosi et al., 1999, 2001; Mukherjee et al., 2001). When plated onto agar containing the lipoprotein stain Congo Red (CR), colonies of *M. avium* segregate into red and white variants (colortypes). Compared to red variants, white variants are more resistant to multiple antibiotics in vitro, more common in patient samples, and better able to survive in mice and human macrophages. The red-white switch is likely to involve the cell surface, where CR binding occurs.

In search of genetic polymorphism between red opaque (RO) and white opaque (WO) variants, we postulated that genetic mutations may be generated by movement of ISs. To test this hypothesis, we adapted the randomly amplified polymorphic DNA (RAPD) technique (Williams et al., 1990) to detect polymorphisms associated with ISs (IS-APD). We identified a novel mobile element, IS*999*, that shows features of the IS*3* family of mobile elements. This report describes the characteristics of IS*999*, including its genetic composition, prevalence in *M. avium* isolates, stability, and the phenotypic consequences of its transposition into structural genes. A unique aspect of this study was the examination of insertional diversity among a large number of clones derived from a single clinical isolate.

2. Materials and methods

2.1. Bacterial isolates and culture conditions

Mycobacterium avium was grown at 37 °C on Middlebrook 7H10 agar (Difco-Becton Dickinson, Franklin Lakes, NJ) containing OADC (oleic acid-albumin-dextrose-catalase) enrichment, 0.5% glycerol and 100 μg/ml CR (MAG-CR agar), prepared as described previously (Cangelosi et al., 1999). In some experiments, cells were grown in Middlebrook 7H9 broth containing ADC enrichment and 0.2% glycerol.

Mycobacterium avium isolates are shown in Table 1. In this report the term 'isolate' refers to the morphotypically heterogeneous population of *M. avium* cells derived from a single patient sample. To minimize the risk of isolating clones with uncharacteristic drug susceptibility profiles, clinical laboratories do not routinely clone *M. avium* isolates as single colonies, however some of the laboratory strains used in this study (strains 102, 104, 501, 2151, and ATCC 35713) had been purified in this manner. We use the term 'clone' to refer to single colonies isolated on the basis of colony type, and their relatively homogenous progeny. RO and WO clones of isolates HMC02, HMC10, HMC04, and 102 were described previously (Cangelosi et al., 1999). The focus of most of the experiments in this report, isolate HMC02, was cultivated from an anonymous patient sample in a Bactec™ system at Seattle's Harborview Medical Center. The contents of the culture were determined to be *M. avium* by AccuProbe™ and IS*1245* analysis (Cangelosi et al., 1999). After four transfers in the clinical lab it was plated directly onto CR agar for isolation of multiple RO and WO clones.

2.2. IS-directed amplified polymorphic DNA (IS-APD) analysis

DNA was obtained by the following procedure for IS-APD and Southern blot analysis. Bacteria were harvested from MAG-CR agar by passing an inoculating loop through multiple morphotypically uniform colonies. Three 10 μl loopfuls were collected into a tube containing 1.2 ml of Breakage buffer (50 mM Tris (pH 7.5), 50 mM KCl, 20 mM $MgCl_2$, 5 mM β-mercaptoethanol) and 0.5 g of 0.1 mm glass beads. The cells were lysed using a bead-beater (BioSpec Products, Inc., Bartlesville, OK) at fast speed for 3 min. Genomic DNA was extracted from the supernatant using phenol and chloroform, precipitated by addition of sodium acetate and isopropanol, washed with 70% ethanol, and dissolved in water. IS-APD reactions were based on the RAPD protocol described previously (Matsiota-Bernard et al., 1997). Genomic DNA (up to 5 ng) was used as template.

Table 1
Mycobacterium avium isolates

Isolate	Description and source	Reference[a]
HMC02	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC04	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC07	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC10	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC15	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC19	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
HMC22	Clinical isolate from Harborview Medical Center, Seattle, WA, USA	A
102	Clinical isolate kindly provided by Clark Inderlied, Children's Hospital, Los Angeles, CA, USA	B
501	Clinical isolate kindly provided by Clark Inderlied, Children's Hospital, Los Angeles, CA, USA	B
104	Clinical isolate kindly provided by William Bishai, Johns Hopkins University; source of DNA for genomic sequencing by The Institute for Genomic Research	C
2151	Laboratory strain kindly provided by Julia Inamine, Colorado State University	D
35713	Reference strain (rough morphotype) purchased from the American Type Culture Collection	–

[a] (A) Cangelosi et al., 1999; (B) Bermudez et al., 1999; (C) Bermudez et al., 1991; (D) Belisle et al., 1993.

Table 2
PCR and IS-APD primers

Name	Sequence	Features (Accession numbers)
Primers used for IS-APD analysis		
P1	5′-GCCGCCGAAACGATCTAC-3′	Sense primer IS*1245*-specific (L33879)
P2	5′-AGGTGGCGTCGAGGAAGAC-3′	Antisense primer IS*1245*-specific (L33879)
DD2	5′-GTCGGGTTGGGCGAAGAT-3′	Sense primer IS1311-specific (AJ223975)
DD3	5′-GTGCAGCTGGTGATCTCTGA-3′	Antisense primer IS1311-specific (AJ223975)
Primers used for standard PCR and Southern blot analysis		
RAP1	5′-GTGGAAGACCGTCGAGGAC-3′	Sense primer IS*999*-specific
RAP2	5′-AGGAACTCAAGATCAGGGATG-3′	Antisense primer hsdR-specific
RAP3	5′-GACGTAGAACTCGCCACCCTC-3′	Sense primer IS*999*-specific
RAP4	5′-TGAAGCGCCCTGGGTTTC-3′	Antisense primer IS*999*-specific
RAP6	5′-CGGTCGATTGCGAAACTC-3′	Antisense primer hsdR-specific
RAP7	5′-ACACGAAGCCGAAACCGGCC-3′	Sense primer hsdR-specific
RAP9	5′-GCGATGAGGATGCCAAGAAG-3′	Sense primer hsdR-specific
LP1	5′-GGCCGCGCTCGAACACCTGAC-3′	Sense primer hdhA-specific
LP2	5′-CGAAATCGCGGGCGTAGTAGA-3′	Antisense primer hdhA-specific
LP3	5′-CGGACCGAGCAAGTCTGGGACGAC-3′	Sense primer hdhA-specific
LP4	5′-AGGCGACCGAGGTGATGTTGATGA-3′	Antisense primer hdhA-specific
Primers used for inverse PCR		
RAP3$^{(-)}$	5′-GAGGGTGGCGAGTTCTACGTC-3′	Complementary to RAP3
RAP4$^{(-)}$	5′-GAAACCCAGGGCGCTTCA-3′	Complementary to RAP4

Primers are listed in Table 2. Amplification products were analyzed by electrophoresis on 2% agarose gels and detected by ethidium bromide staining.

2.3. Southern blot analysis

Up to 5 μg of genomic DNA, prepared as described above, was digested to completion with 10 units of restriction endonuclease *Sac*II (New England Biolabs, Inc., Beverly, MA). After 0.7% agarose gel electrophoresis, DNA was partially hydrolyzed and transferred onto Nytran membranes. Probes were prepared and labeled during PCR from *M. avium* DNA by incorporation of Biotin-dCTP using the PCR DNA biotinylation kit (KPL, Inc., Gaithersburg, MD). The following pairs of primers were used to prepare probes: RAP3-RAP4 specific to IS*999*, P1-P2 specific to IS*1245*, RAP6-RAP7 specific to *hsd*R, LP1-LP2 specific to *hdhA* (Table 2). Hybridization and detection were performed as described by the kit manufacturer.

2.4. Statistical analysis

The χ^2 test was used to test the correlation between colortype and the presence of IS*999* in specific genes. The null hypothesis for each test was that IS*999* insertion into a gene is randomly distributed between colortypes (error risk 0.5%).

2.5. Molecular tools using the internet

We used the Basic Local Alignment Search Tool (BLAST) and the ORF Finder tool at the NCBI web site (http://www.ncbi.nlm.nih.gov/). Motifs in protein sequences were identified by using Blocks, which detects local regions of similarity (http://www.blocks.fhcrc.org/), and ProDom, which detects end-to-end similarity (http://protein.Toulouse.inra.fr/prodom/blast_form.html). When needed, *Cognitor* (http://www.ncbi.nlm.nih.gov/) was used to corroborate results obtained by BLAST. It searches for clusters of orthologous groups (COGs) which are clusters of very similar proteins found in at least three species. We are grateful to The Institute for Genomic Research (www.tigr.org) for making available the preliminary, unannotated *M. avium* genome sequence. The nucleotide sequence of the genomic region of isolate HMC02 containing IS*999* and a portion of *hsdR* has been deposited in GenBank (Accession number: AF232829).

3. Results

3.1. IS-APD analysis of WO and RO variants

This study began with the hypothesis that unidentified ISs contribute to colortype variation in *M. avium*. To test this hypothesis, we took advantage of the common organization of mycobacterial ISs, most of which carry two imperfect inverted repeats (IRs) encompassing one or more open reading frame (ORF) coding for transposases. ISs are grouped according to conserved transposase gene sequences. IS-APD primers were designed to target such conserved sequences, allowing the primers to anneal to multiple ISs during low stringency PCR. When an IS primer also found an imperfect matching sequence in genomic DNA adjacent to an IS, an amplification product was generated. Electro-

phoretic analysis of such products yielded fingerprints associated with IS insertion patterns. We chose four primers complementary to elements of the two most abundant IS families in the *M. tuberculosis* genome (Cole et al., 1998), IS*3* and IS*256* (Table 2). With all of these primers but P2, IS-APD patterns did not differ between representative RO and WO clones of *M. avium* isolates HMC02, HMC10 and 102 (data not shown). With primer P2, derived from IS*1245* (Matsiota-Bernard et al., 1997), we consistently observed a 660 bp band amplified from the genome of a representative WO clone of isolate HMC02, but not from the genome of a representative RO clone of the same isolate (Fig. 1). This band was not visible in the IS-APD profiles of RO and WO clones of two other *M. avium* isolates, HMC10 and 102.

The 660 bp IS-APD band was purified from an agarose gel. Nucleotide sequence analysis of the fragment revealed two fused regions. One region of 194 bp was 82% homologous to IS*1601* of *M. avium* (Eckstein et al., 2000), and the other region formed a partial ORF. The complete nucleotide sequence of the genomic region surrounding this ORF was obtained by inverse PCR using primers directed outwardly from the 194 bp IS-like region ($RAP3^{(-)}$-$RAP4^{(-)}$, Table 2), combined with genomic cloning and automated DNA sequencing. It codes for a putative protein of 877 amino acids. The *Cognitor* program placed this protein in a family of restriction enzyme helicase subunits (COG0610). It shares 45% amino acid homology with the *hsdR* gene of *Escherichia coli*, which codes for the R subunit of a Type I restriction-modification system. Immediately downstream of *hsdR* was another ORF, *hsdM*, coding for a putative protein with 61% amino acid homology to the M subunit of the same system in *E. coli*. The *hsdM* ORF had IS*999* insertions in the RO as well as WO clones.

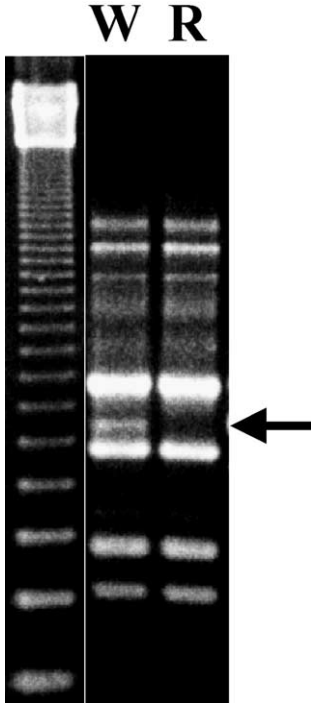

Fig. 1. IS-APD analysis of representative WO (lane W) and RO (lane R) variants of *M. avium* isolate HMC02. The IS-APD primer was P2. The position of the 660 bp band specific to the white variant is indicated by the arrow.

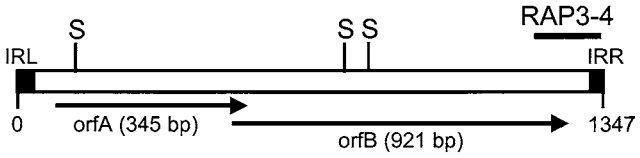

Fig. 2. Diagram of IS*999*. *Sac*II sites (S) and positions of the two overlapping ORFs are indicated. Left (IRL) and right (IRR) inverted repeats are represented by black boxes. The region amplified by PCR primers RAP3 and RAP4 to generate the IS*999*-specific probe is labeled RAP3-4.

3.2. Characterization of IS999

Sequence analysis of the complete IS-like region in *hsdR* revealed a previously unknown 1347 bp element which we named IS*999* (Fig. 2). The following features of IS*999*, detailed in GenBank Accession number AF232829, relate it to the IS*51* group of the IS*3* family (Mahillon and Chandler, 1998). It has a high G + C content (64%) similar to the rest of the genome. It has 29 bp IRs and is terminated by 5′-TG…CA-3′. The 5′ IR (IRL) and the 3′ IR (IRR) contain blocks of cytosine and guanosine residues and are respectively 75% and 86% homologous to the IRs of IS*51*. Upon insertion into *hsd*R, IS*999* produced 3 bp direct repeats (5′-ATC-3′). IS*999* appears to possess two overlapping ORFs (orfA and orfB), which are respectively 64% and 65% homologous at the amino acid level to orfA and orfB of IS*987*, and 93% and 85% homologous to orf3 and orf4 of IS1601 (GenBank Accession numbers: X57835 and AFO60182). In IS*999*, the 344 bp orfA and the 917 bp orfB are in the relative translational reading frames 0 and −1, respectively. This is analogous to IS*987*, however the consensus translational frameshift signal (5′-XXXYYYZ-3′) was not found in IS*999*. The two ORFs contain motifs typical of IS*3* transposases: a helix-turn-helix in orfA from residue 31 to 70, a DD(35)E motif in orfB from residue 137 to 266 and a potential leucine zipper in orfA from residue 287 to 372 (Mahillon and Chandler, 1998).

Southern blot analysis of *M. avium* isolates HMC02, HMC10, 102, and 104 showed that IS*999* is present in copy numbers ranging from four to 16 (Fig. 3A). From the standpoint of strain resolution, IS*999* analysis compares favorably to IS*1245* analysis (Fig. 3B). However, IS*999* was not detected in all *M. avium* isolates examined. By PCR and Southern blot analysis, it was detected in ten isolates (HMC02, HMC07, HMC10, HMC15, HMC19, HMC22, 102, 104, 501, 2151) but was absent in two other isolates (HMC04 and ATCC 35713).

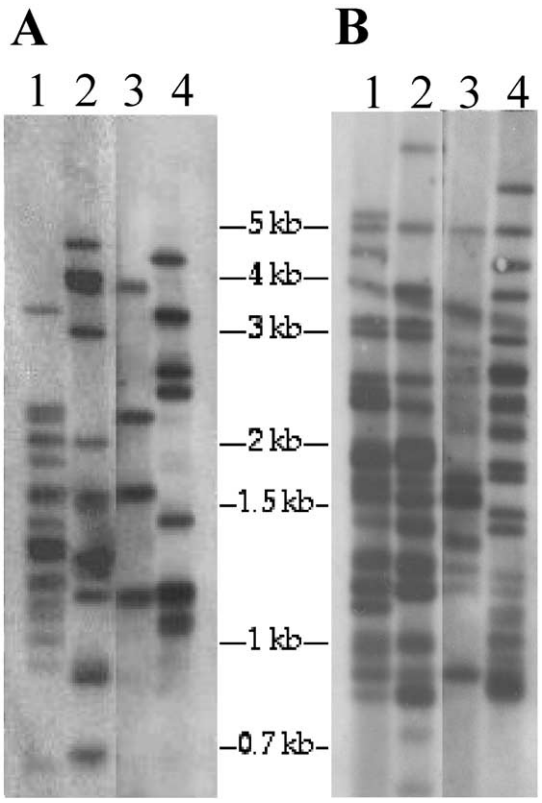

Fig. 3. RFLP analysis using probes specific for IS*999* (A) and IS*1245* (B). DNA extracted from *M. avium* isolates 104 (1), HMC10 (2), 102 (3), and HMC02 (4) was digested by *Sac*II, separated on a 1.1% agarose gel, and transferred onto a nylon membrane for Southern hybridization analysis. Probes specific for IS*999* and IS*1245* were prepared by PCR using primer sets RAP3-RAP4 and P1-P2, respectively (Table 2).

3.3. Characterization of IS999 stability in red and white variants

RFLP analysis with IS*999*- and IS*1245*-specific probes was used to assess the stability of IS*999* within a single isolate, HMC02. We also asked whether IS*999* mobility was related to red-white morphotypic variation, as suggested by IS-APD results. The archived HMC02 sample derived from an anonymous patient was plated to obtain 29 RO colonies and 31 WO colonies as described in Section 2. RFLP analysis of the 60 clones revealed eight different IS*999* banding patterns that fell into two groups, A1 through A5 (A group) and B1 through B3 (B group) (Fig. 4A). Representative clones exhibiting each of the eight IS*999* patterns were also examined by IS*1245* RFLP. All clones in the A group shared a single IS*1245* pattern, while all clones in the B group shared a different IS*1245* pattern (Fig. 4B). Polyclonal infections with *M. avium* have been reported elsewhere (Arbeit et al., 1993; Slutsky et al., 1994; Wallace et al., 1998). The B group consisted entirely of WO clones, while the A group contained both colortypes.

Southern hybridization analysis with a PCR-generated *hsdR* probe (RAP6-RAP7) showed that the polymorphism labeled U in Fig. 4A corresponded to the IS*999* insertion into *hsd*R discovered by IS-APD. The polymorphism labeled L, a fragment of approximately 1000 bp, was cloned and sequenced by inverse PCR as described above. The cloned fragment consisted of IS*999* fused to a 462 bp sequence that was 97% identical to an ORF in the unannotated preliminary genome sequence of *M. avium* strain 104 (www.tigr.org). This ORF codes for a putative protein of 340 amino acids that shares 42% amino acid homology with the *E. coli* 7-a-hydroxysteroid dehydrogenase (HdhA, EC 1.1.1.159). The *Cognitor* program placed this ORF in the family of short-chain alcohol dehydrogenases (SCAD) involved in fatty acid metabolism (COG1028).

The L polymorphism was not randomly distributed between the 29 RO colonies and 31 WO colonies ($P = 0.005$). Of the 27 clones with IS*999* inserted into *hdhA*, 23 (85.2%) were RO. Of the 33 clones with no IS999 insertion at this site, 25 (75.7%) were WO. The correlation was not absolute, so it did not point to a direct role for *hdhA* in colortype. However, the apparent bias in the distribution of IS*999* insertion at this site can be explained by phenotypic effects that are tolerated differently in the red and white morphotypic backgrounds. None of the other polymorphisms visible in Fig. 4 showed a significant correlation with colortype.

3.4. IS999 insertion sites

To determine whether this unstable element might affect the functions of other genes, we characterized IS*999* insertion sites in three *M. avium* isolates. An inverse PCR approach was used to clone and sequence insertional junctions. This strategy yielded sequences of eight insertion sites from isolate HMC02, three from isolate 102, and three from isolate HMC10. Of the 14 sites analyzed, 12 were in likely protein-coding genes identified by using ORF Finder and BLAST (Table 3).

The incomplete genome sequence of *M. avium* isolate 104 (www.tigr.org) was also examined. At the time of this writing, the genome was assembled into a single unannotated 5.48 megabase contig, allowing us to assess the distribution of IS999 insertions around the genome. Nine complete copies and one 386 bp fragment of IS999 were found. Seven of these ten insertion sites were within likely protein-coding genes identified by using ORF Finder and BLAST (Table 4). All ten were situated between positions 1,661,546 and 3,414,095, a region encompassing about a third of the contig. Three sites were clustered in a 55 kb region between positions 3,359,078 and 3,414,095. This region has strong homology to the ser2 gene cluster that codes for cell surface glycopeptidolipid synthesis (Eckstein et al., 2000). The ser2 cluster is a hotspot for IS insertion in *M. avium* strain 2151 (GenBank Accession number: AF143772). Another three sites were clustered in a 101 kb region between positions 2,167,551 and 2,268,378. Thus,

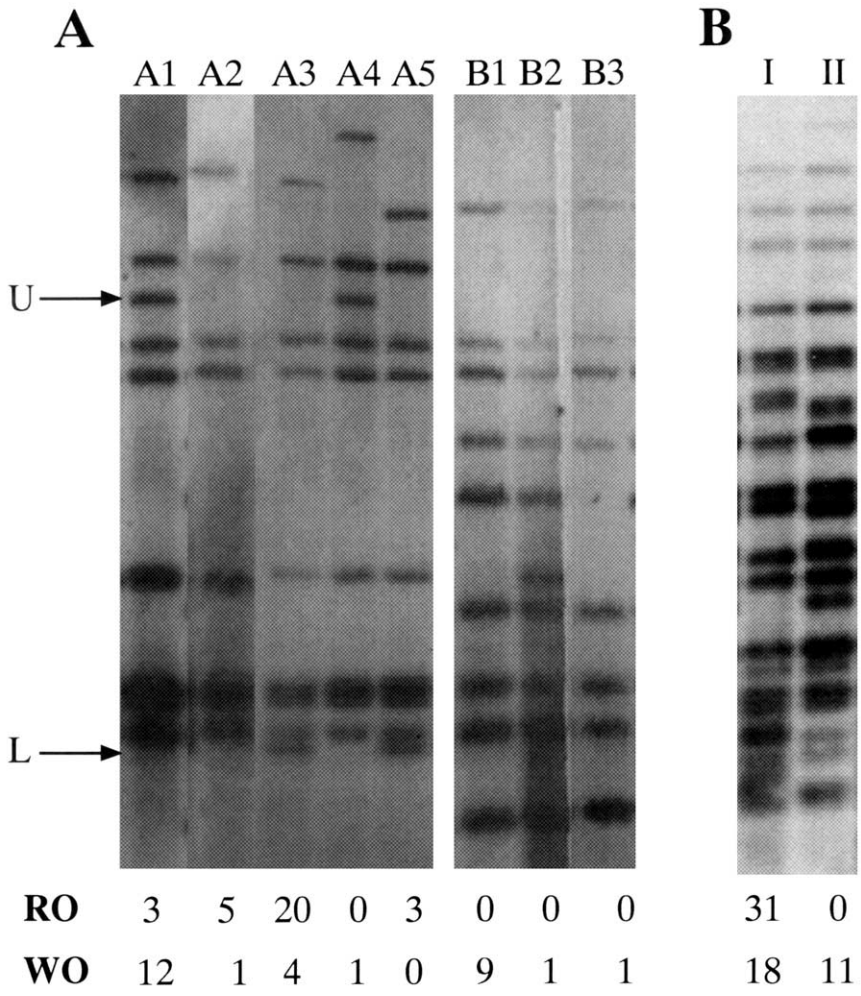

Fig. 4. RFLP patterns of 60 clones of *M. avium* isolate HMC02. The number of RO and WO clones exhibiting each pattern is shown in the box below. DNA was digested by *Sac*II, separated on a 1.1% agarose gel, transferred onto nylon membrane, and hybridized with probes specific for IS*999* (A) or IS*1245* (B). Probes were prepared as in Fig. 3. Clones exhibiting the A1, A2, A3, A4, and A5 patterns in (A) all had the same IS*1245* profile, labeled I in (B). Clones exhibiting the B1, B2, and B3 patterns in (A) all had a different IS*1245* profile, labeled II in (B).

six of the ten IS999 insertion sites were concentrated in two regions that together constitute less than 3% of the genome. While the map may change with final editing of the genome sequence, these numbers strongly suggest a nonrandom distribution of IS999 insertion sites in the *M. avium* strain 104 genome.

Two additional observations are worth noting. First, two IS*999* insertion sites in isolate 104 are identical to insertion sites in isolate HMC10 (Table 3). One is in a gene with homology to the *lpr*M gene of *M. tuberculosis*. Isolates 104 and HMC10 both have IS*999* inserted into the sequence CGCCGGCCGGGAAGC-IS*999*-AGCCGACATGCGCAG within this gene (the underlined bases are direct repeats found at most IS*999* insertion sites). The second was in an ORF coding for a putative glycerol-3-phosphate acyltransferase. Both isolates have IS*999* inserted into the sequence TTGCAGCAGAATGTC-IS*999*-GTCTTCGCTGCGGCC at this site. Although the two isolates are not epidemiologically linked, they have somewhat similar IS*1245* insertion patterns (Fig. 3B). Therefore, the shared IS*999* insertion sites may be remnants of a common ancestor.

The second noteworthy observation was the insertion of IS999 into the transposase genes of two copies of IS1245 (Table 4). Similar arrangements are visible in Table 3, with IS*999* inserted into the transposase genes of transposable elements IS*1601*-A and ID*780*. One expected outcome of such nested and clustered insertions is that transposition of one element would affect RFLP patterns of others. Such a mechanism might help explain the correlation between IS*999* and IS*1245* patterns within isolate HMC02.

4. Discussion

The IS-APD technique proved useful for detecting polymorphisms associated with previously unknown ISs in *M. avium*. It could be adapted for similar use in any microbial genome for which conserved IS sequences are known.

Table 3
Insertion loci sequenced from isolates HMC02, HMC10, and 102

Isolate	Insertion site: description of closest homolog (% amino acid homology)	Species and GenBank Accession number of closest homolog
HMC02	Transport membrane protein gene *mmpl12* (63%)	*Mycobacterium tuberculosis* – Q50585
HMC02	Type I restriction-modification R subunit gene *hsdR* (41%)	*Escherichia coli* – P08956
HMC02	Type I restriction-modification M subunit gene *hsdM* (61%)	*Escherichia coli* – AAA23985.1
HMC02	Transmembrane protein gene M1CB 1886.05c (55%)	*Mycobacterium leprae* – CAA22902.1
HMC02	ID780 transposase gene (51%)	*Bradyrhizobium japonicum* – AAG61025
HMC02	Ribitol dehydrogenase gene C45B11.3 (62%)	*Caenorhabditis elegans* – T19954
HMC02	7a-hydroxysteroid dehydrogenase gene *hdhA* (42%)	*Escherichia coli* – P25529
HMC02	Apparent intergenic region	–
102	IS1601-A transposase gene (100%)	*Mycobacterium avium* – AAC71696.1
102	Hypothetical protein Rv3787c (63%)	*Mycobacterium tuberculosis* – G70696
102	Cytochrome P450 gene Rv1785c (91%)	*Mycobacterium tuberculosis* – C70929
HMC10	Lipoprotein gene *lprM* (60%)[a]	*Mycobacterium tuberculosis* – NP_216486
HMC10	Glycerol-3-phosphate acyltransferase gene (100%)[a]	*Mycobacterium tuberculosis* – CAA16059
HMC10	Apparent intergenic region	–

[a] Identical to IS*999* insertion sites in isolate 104 genome sequence.

When applied to *M. avium*, IS-APD revealed a new mobile element, IS*999*. A unique aspect of this study was the examination of 60 individual clones of a single patient isolate. Within this collection of clones, eight IS*999* RFLP groups were observed within two IS*1245* RFLP groups, suggesting that IS999 is a relatively less stable element.

In addition to their known role in genetic variation, ISs may also promote phenotypic variation. This was shown to be true of some bacterial pathogens (Ziebuhr et al., 1999; Simonet et al., 1996), but reports of such activity in mycobacteria are rare. In the *M. tuberculosis* H37Rv genome, most IS insertions are in intergenic regions (Cole et al., 1998). However, one study found that >63% of IS*6110* insertion sites were in genuine coding regions in *M. tuberculosis* (Sampson et al., 1999). Our analysis of IS*999* insertion sites in four isolates of *M. avium* also found a high incidence of insertion into putative genes. If one assumes that at least some of these ORFs are true genes, then transposition of IS*999* into and out of them would be expected to contribute to phenotypic variation. ISs can also drive phenotypic variation in *M. avium* by providing regions of homology for deletions mediated by homologous recombination (Eckstein et al., 2000).

A correlation was observed between IS*999* insertion into the *hdhA* and the red colortype ($P = 0.005$). The putative *hdhA* gene product belongs to a family of dehydrogenases that function in fatty acid metabolism. The imperfect nature of the correlation is not consistent with a direct role for *hdhA* in colortype. The simplest model to explain it is that insertional inactivation of *hdhA* affects fatty acid metabolism in a manner that, for unknown reasons, is less detrimental to red cells than to white cells.

IS999 was first identified within a putative gene, *hsd*R, that codes for a product that is strongly homologous to the R subunit of a Type I restriction-modification system. A gene coding for a protein homologous to the M subunit of the same system (*hsdM*) is immediately downstream in isolate HMC02. Restriction-modification activity has been reported in *M. avium* (Crawford et al., 1981), but this is the first description of Type I restriction-modification genes in this pathogen. The *hsd* region was not found in the unassembled genome sequence of strain 104 (www.tigr.org), and its

Table 4
Insertion loci in the assembled genome sequence of isolate 104

Position (base number)	Insertion site: description of closest homolog (% amino acid homology)	Species and GenBank Accession number of closest homolog
1,661,546	Glycerol-3-phosphate acyltransferase *plsB2* (84%)[a]	*Mycobacterium tuberculosis* – CAA16059
1,880,596	Apparent intergenic region	–
2,167,551	IS1245 transposase gene (100%)	*Mycobacterium avium* – AAA69904
2,267,032	Conserved hypothetical protein (40%)	*Agrobacterium tumefaciens* – AAK90396
3,405,056	Apparent intergenic region within the ser2 cluster (98%)	*Mycobacterium avium* – AF143772
2,063,674	Lipoprotein *lprM* (66%)[a]	*Mycobacterium tuberculosis* – NP_216486
2,224,076	386-base fragment of IS999 inserted into apparent intergenic region	–
2,729,860	IS1245 transposase gene (100%)	*Mycobacterium avium* – AAA69904
3,412,749	Bacteriophage A-protein-like gene within the ser2 cluster (100%)	*Mycobacterium avium* – AAD44204
3,359,078	Peptide synthetase gene within the ser2 cluster (100%)	*Mycobacterium avium* – AAF63833

[a] Identical to IS*999* insertion sites in isolate HMC10.

absence in that strain was confirmed by PCR and Southern analysis (data not shown).

In summary, IS*999* is a mobile element found in ten out of 12 *M. avium* isolates examined. A striking characteristic of *M. avium* is its ability to thrive in diverse environments ranging from fresh water to the macrophages of humans and other animals. The completion of an *M. avium* genome sequence will make it possible to use functional and comparative genomic approaches to characterize the mechanisms by which this adaptation occurs. However, such approaches will be complicated by the high degree of mutability this organism exhibits, even within single strains defined by pedigree, epidemiology, or IS*1245* RFLP (Arbeit et al., 1993; Matsiota-Bernard et al., 2000; von Reyn et al., 1995). IS*999* appears to have better substrain discrimatory power than IS*1245*, making it a potentially useful marker for tracking the genetic drift that *M. avium* isolates undergo in the laboratory and in nature.

Acknowledgements

We wish to thank Kambiz Yaraei for his help in characterizing clones of *M. avium* isolate HMC02. This work was supported by grant AI25767 from The National Institutes of Health and grant G8E10521 from The U.S. Environmental Protection Agency. J.-P.L. was supported in part by NIH Training Grant #T32AI07509. Preliminary *M. avium* genome sequence data were obtained from The Institute for Genomic Research website at http://www.tigr.org. Sequencing of the *M. avium* genome is being carried out by The Institute for Genomic Research with support from the National Institutes of Health.

References

Arbeit, R.D., Slutsky, A., Barber, T.W., Maslow, J.N., Niemczyk, S., Falkinham 3rd, J.O., O'Connor, G.T., von Reyn, C.F., 1993. Genetic diversity among strains of *Mycobacterium avium* causing monoclonal and polyclonal bacteremia in patients with AIDS. J. Infect. Dis. 167, 1384–1390.

Bauer, J., Andersen, A.B., Askgaard, D., Biese, S.B., Larsen, B., 1999. Typing of clinical *Mycobacterium avium* complex strains cultured during a 2-year period in Denmark by using IS1245. J. Clin. Microbiol. 37, 600–605.

Belisle, J.T., Brennan, P.J., 1994. Molecular basis of colony morphology in *Mycobacterium avium*. Res. Microbiol. 145, 237–242.

Belisle, J.T., McNeil, M.R., Chatterjee, D., Inamine, J.M., Brennan, P.J., 1993. Expression of the core lipopeptide of the glycolipid surface antigens in rough mutants of *Mycobacterium avium*. J. Biol. Chem. 268, 10510–10516.

Bermudez, L.E., Petrofsky, M., Kolonoski, P., Young, L.S., 1991. An animal model of *Mycobacterium avium* complex disseminated infection after colonization of the intestinal tract. J. Infect. Dis. 165, 75.

Bermudez, L.E., Kolonoski, P., Wu, M., Aralar, P.A., Inderlied, C.B., Young, L.S., 1999. Mefloquine is active in vitro and in vivo against the *Mycobacterium avium* complex. Antimicrob. Agents Chemother. 43, 1870–1874.

Campbell, A., 1981. Evolutionary significance of accessory DNA elements in bacteria. Annu. Rev. Microbiol. 35, 55–83.

Cangelosi, G.A., Palermo, C.O., Laurent, J.-P., Hamlin, A.M., Brabant, W.H., 1999. Colony morphotypes on Congo Red agar segregate along species and drug susceptibility lines in the *Mycobacterium avium-intracellulare* complex. Microbiology 145, 1317–1324.

Cangelosi, G.A., Palermo, C.O., Bermudez, L.E., 2001. Phenotypic consequences of red-white colony type variation in *Mycobacterium avium*. Microbiology 147, 527–533.

Cole, S.T., et al., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393, 537–544.

Crawford, J.T., Cave, M.D., Bates, J.H., 1981. Evidence for plasmid-mediated restriction-modification in *Mycobacterium avium intracellulare*. J. Gen. Microbiol. 127, 333–338.

Eckstein, T.M., Inamine, J.M., Lambert, M.L., Belisle, J.T., 2000. A genetic mechanism for deletion of the *ser*2 gene cluster and formation of rough morphological variants of *Mycobacterium avium*. J. Bacteriol. 182, 6177–6182.

Fang, Z., Doig, C., Kenna, D.T., Smittipat, N., Palittapongarnpim, P., Watt, B., Forbes, K.J., 1999. IS*6110*-mediated deletions of wild-type chromosomes of *Mycobacterium tuberculosis*. J. Bacteriol. 181, 1014–1020.

Gaffney, T.D., Lessie, T.G., 1987. Insertion-sequence-dependent rearrangements of *Pseudomonas cepacia* plasmid pTGL1. J. Bacteriol. 169, 224–230.

Gordon, S.V., Heym, B., Parkhill, J., Barrell, B., Cole, S.T., 1999. New insertion sequences and novel repeated sequence in the genome of *Mycobacterium tuberculosis* H37Rv. Microbiology 145, 881–892.

Hernandez-Alles, S., Benedi, V.J., Martinez-Martinez, L., Pascual, A., Aguilar, A., Tomas, J.M., Alberti, S., 1999. Development of resistance during antimicrobial therapy caused by insertion sequence interruption of porin genes. Antimicrob. Agents Chemother. 43, 937–939.

Mahillon, J., Chandler, M., 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62, 725–774.

Matsiota-Bernard, P., Waser, S., Tassios, P.T., Kyriakopoulos, A., Legakis, N.J., 1997. Rapid discrimination of *Mycobacterium avium* strains from AIDS patients by randomly amplified polymorphic DNA analysis. J. Clin. Microbiol. 35, 1585–1588.

Matsiota-Bernard, P., Zinzendorf, N., Onody, C., Guenounou, M., 2000. Comparison of clarithromycin-sensitive and clarithromycin-resistant *Mycobacterium avium* strains isolated from AIDS patients during therapy regimens including clarithromycin. J. Infect. 40, 49–54.

Mukherjee, S., Petrofsky, M., Yaraei, K., Bermudez, L.E., Cangelosi, G.A., 2001. The white morphotype of *Mycobacterium avium-intracellulare* is common in infected humans and virulent in infection models. J. Infect. Dis. 184, 1480–1484.

Sampson, S.L., Warren, R.M., Richardson, M., van der Spuy, G.D., van Helden, P.D., 1999. Disruption of coding regions by IS*6110* insertion in *Mycobacterium tuberculosis*. Tuber. Lung Dis. 79, 349–359.

Simonet, M., Riot, B., Fortineau, N., Berche, P., 1996. Invasin production by *Yersinia pestis* is abolished by insertion of an IS*200*-like element within the inv gene. Infect. Immun. 64, 375–379.

Slutsky, A.M., Arbeit, R.D., Barber, T.W., Rich, J., von Reyn, C.F., Pieciak, W., Barlow, M.A., Maslow, J.N., 1994. Polyclonal infections due to *Mycobacterium avium* complex in patients with AIDS detected by pulsed-field gel electrophoresis of sequential clinical isolates. J. Clin. Microbiol. 32, 1773–1778.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S., Musser, J.M., 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. USA 94, 9869–9874.

van Embden, J.D.A., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., Hermans, P., Martin, C., McAdam, R., Shinnick, T.M., Small, P.M., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. 31, 406–409.

van Soolingen, D., Bauer, J., Ritacco, V., Cardoso Leao, S., Pavlik, I., Vincent, V., Rastogi, N., Gori, A., Bodmer, T., Garzelli, C., Garcia,

M.J., 1998. IS*1245* restriction fragment length polymorphism typing of *Mycobacterium avium* isolates: proposal for standardization. J. Clin. Microbiol. 36, 3051–3054.

von Reyn, C.F., Jacombs, N.J., Arbeit, R.D., Maslow, J.N., Niemczyk, S., 1995. Polyclonal *Mycobacterium avium* infections in patients with AIDS: variations in antimicrobial susceptibilities of different strains of *M. avium* isolated from the same patient. J. Clin. Microbiol. 33, 1008–1010.

Wallace Jr., R.J., Zhang, Y., Brown, B.A., Dawson, D., Murphy, D.T., Wilson, R., Griffith, D.E., 1998. Polyclonal *Mycobacterium avium* complex infections in patients with nodular bronchiectasis. Am. J. Respir. Crit. Care Med. 158, 1235–1244.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535.

Ziebuhr, W., Krimmer, V., Rachid, S., Lossner, I., Gotz, F., Hacker, J., 1999. A novel mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS*256*. Mol. Microbiol. 32, 345–356.